DTU Library

Applying Multi-Class Support Vector Machines for performance assessment of shipping operations: The case of tanker vessels

Pagoropoulos, Aris; Møller , Anders H.; McAloone, Tim C.

Published in:
Ocean Engineering

Link to article, DOI:
10.1016/j.oceaneng.2017.05.001

Publication date:
2017

Document Version
Peer reviewed version

Citation (APA):
Pagoropoulos, A., Møller , A. H., & McAloone, T. C. (2017). Applying Multi-Class Support Vector Machines for performance assessment of shipping operations: The case of tanker vessels. *Ocean Engineering*, *140*, 1-6. https://doi.org/10.1016/j.oceaneng.2017.05.001

Article for Journal of Ocean Engineering

Title: Applying Multi-Class Support Vector Machines for performance assessment of shipping operations: The case of tanker vessels

Authors: Aris Pagoropoulos[1*], Anders H. Møller[2], Tim C. McAloone[1]

[1] Technical University of Denmark, Department of Mechanical Engineering, Nils Koppels Allé, Building 426, 2800 Kgs. Lyngby, Denmark. E-mails: arispa@mek.dtu.dk, tcma@dtu.dk

[2] TORM A/S, Tuborg Havnevej 18, 2900 Hellerup, Denmark. E-mail: ahm@torm.com

Manuscript Correspondence:

*Corresponding author: Aris Pagoropoulos

Address: Nils Koppels Allé, Building 426

Email: arispa@mek.dtu.dk

Telephone: +45 27642264

Abstract

Energy efficient operations are a key competitive advantage for modern shipping companies. During the operation of the vessel, improvements in energy use can be achieved by not only by technical upgrades, but also through behavioural changes in the way the crew on board is operating the vessels. Identifying the potential of behavioural savings can be challenging, due to the inherent difficulty in analysing the data and operationalizing energy efficiency within the dynamic operating environment of the vessels. This article proposes a supervised learning model for identifying the presence of energy efficient operations. Positive and negative patterns of energy efficient operations were identified and verified through discussions with senior officers and technical superintendents. Based on this data, the high dimensional parameter space that describes vessel operations was first reduced by means of feature selection algorithms. Afterwards, a model based on Multi- Class Support Vector Machines (SVM) was constructed and the efficacy of the approach is shown through the application of a test set. The results demonstrate the importance and benefits of machine learning algorithms in driving energy efficiency on board, as well as the impact of power management on energy costs throughout the life cycle of the ships.

Keywords: tankers; energy efficiency; machine learning; support vector machines

1 Introduction

There are strong economic and environmental incentives in reducing the fuel consumption of the shipping industry. The need to curb the increase in the global average temperatures , together with the designation of new emission control areas in China underline the importance of energy management on board modern vessels.

Interestingly, within energy management systems, shipping has attracted limited attention. In a recent review by Lee & Cheng (2016) the authors argue that although energy management systems have been extensively studied for over 40 years, the majority of studies are focused on either buildings or industrial and factory energy management systems with no studies on shipping. In the shipping literature a number of works have attempted to develop models that simulate the performance of the ship energy systems and identify energy consumption patterns. Trodden et al. (2015) propose a data analysis methodology to isolate the steady-state free-running condition of a harbour tug. The developed algorithm separates the data-stream, as output from monitoring devices, into periods associated with steady-state, free-running condition, and non-steady-state free-running condition and shows that the tug is being operated in a fuel efficient manner, making the most of a retrofitted economy engine speed selector. Cichowitz et al. (2015) discuss the use of Dynamic Energy Modelling (DEM) for realistic simulation of ship energy systems. DEM captures holistically the transfer, conversion and storage of energy on board a ship as a function of its operational profile and over long periods of time or during its commercial life-cycle. Simulation using DES is presented for four hypothetical scenarios that illustrate the feasible operational space for the case of a container ship. Similar studies can be found on other industrial sectors such as household equipment (Murray et al. 2016) and hybrid vehicles (He et al. 2016).

All studies that were just described acknowledge the growing importance of data and data analysis, and their potential in operationalizing performance management across the shipping industry. In a study of the digital transformation conducted by the MIT centre for digital business, Westerman et al. (2011) argue that performance management is one of the building blocks of the ongoing digital transformation. In the oil and gas industry, DNV – GL claim that if the oil and gas industry could analyse and understand all the data it is currently producing in a more coordinated manner, operational efficiency could be boosted by as much as 20%. However the same report warns that the potential of big data is hampered by a lack of resources, lack of experience and the increasing volume of data (DNV - GL 2016).

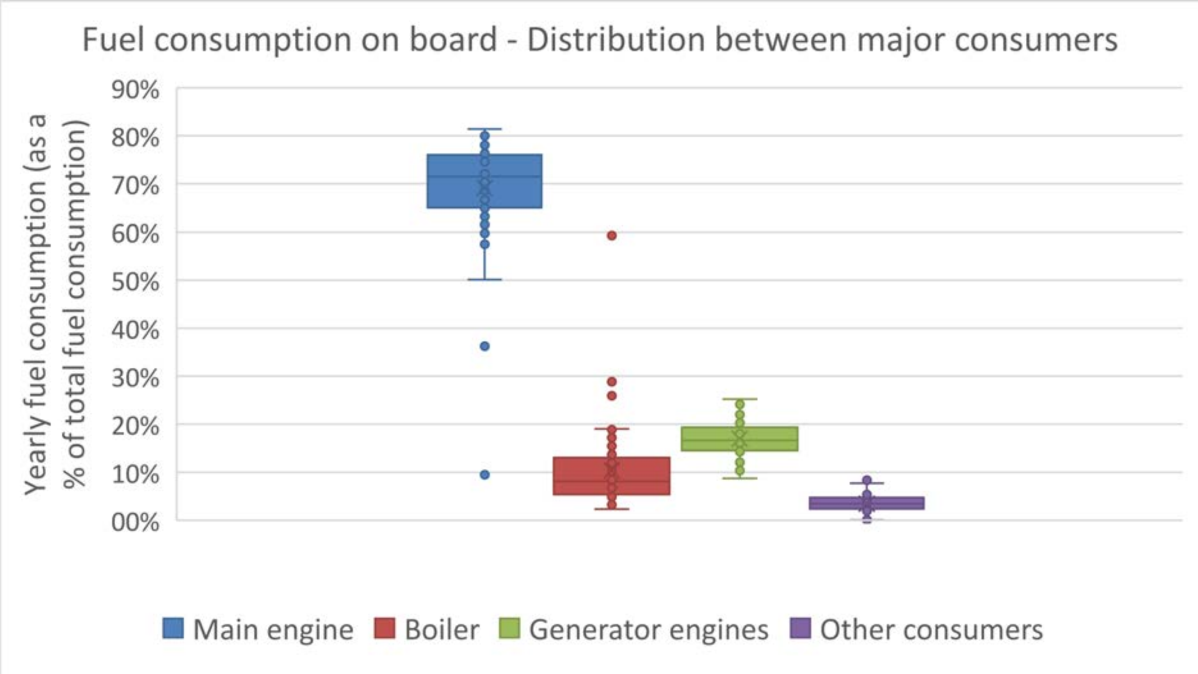
However, simply measuring fuel consumption is not enough in driving energy efficiency. Trodden et al. (2015) argue that while data monitoring devices are relatively inexpensive, the process of analysing data can be complex, particularly when a ship's activities are diverse. In their study of the German and Danish shipping industry, Poulsen and Johnson (2015) conclude that the lack of information on energy efficiency and lack of time to produce and provide reliable energy efficiency information cause energy efficiency gaps.

Data-related challenges are not confined to shipping. A recent analysis from the McKinsey Global Institute argues that even in established organizations where core processes are centred around data analytics, management-approval processes have not kept up with the advancements in data analytics (Court 2015). However, the shipping industry exhibits certain characteristics that further complicate data analysis. The different characteristics of power generation systems and consumers for vessels in operation require careful consideration and adjustment of energy consumption profiles to ship-specific characteristics. Especially in the tramp shipping market that is driven by the complex balance of supply and demand (Stopford 2009), operating profiles can change rapidly. Energy consumption patterns are also influenced by safety considerations. For example, specific equipment according to the ship safety plan might be turned on when transiting high risk areas (NATO Shipping Centre 2016). But most importantly, as vessels engage in a multitude of operational activities, energy consumption patterns need to be associated to those particular activities (Trodden et al. 2015).

Challenges towards data analysis can also stem from the various ship management models that appear in the shipping industry. Information and incentives are often fragmented, as fuel consumption is a performance measure of the commercial department – reflected in the Time Charter Equivalent (TCE) - and often outside of the sphere of influence of the technical department, which is often primarily measured on Operational Expenses (OPEX). This paradox is even more prevalent in third party ship management, where information regarding fuel saving potential is not readily provided and shared by decision makers at sea and ashore (Poulsen & Sornn-Friese 2015). This situation creates data silos in shipping companies and invites for one of the key business challenges of the modern age according to Thompson (2012), which is to recognize and use the valuable information that is scattered around the organization.

In light of the above mentioned challenges, we believe that these limitations can be overcome by a methodological shift to multivariate machine learning techniques. To the best of our knowledge, machine learning techniques have not been applied on power management in shipping – at least not in the open literature. However, machine learning has been extensively used for power management in other industrial sectors, with particular focus on prediction. In a recent review of forecasting approaches for the building sector, Chalal et al. (2016) argue that Support Vector Machines (SVM) and artificial neural networks models (ANN) are the most common tools, to develop energy prediction approaches, which in turn support physical improvement strategies. Especially SVMs have been used for time series predictions, particularly in financial time series and electrical load forecasting (Sapankevych & Sankar 2009).

Given their wide adoption and alleged benefits, we investigated the efficacy of Support Vector Machines in eliciting the correct information from the energy consumption patterns. Based on the results of that analysis, we assess the potential savings from behavioural improvements. This article proposes a supervised learning model for identifying the presence of energy efficient operations, as a basis for developing an energy management methodology. Focus is on the production of electricity on board a group of tanker vessels. Production of electricity on board from generator engines comprises between 9 % and 25% of the total fuel consumption of a tanker vessel (Figure 1). Through ship-specific adjustments, the proposed methodology evaluates operational practices between different vessels, thus providing an informed picture of the behaviour-driven efficiency on-board. The performance and accuracy of the classifier was evaluated by means of 5-fold cross validation. The development and scope of the methodology, while novel in the shipping literature, follows extant directions for future research to identify actual effects of fuel initiatives (Poulsen & Johnson 2015), measured under comparable conditions (Trodden et al. 2015).



z

Figure 1: Tanker fuel consumption distribution per major consumer

2 Description of the proposed methodology

2.1 Energy efficiency

As discussed in Section 1, electricity production on board is influenced by multiple factors, and therefore consumption patterns need to be associated to the operational conditions of the vessel and analysed under comparable conditions. When looking at the typical operational profile of a vessel, operations such as loading of cargo or sailing are characterized by rather steady and predictable consumption patterns. In these cases the vessels are mostly idle and several systems such as the engine cooling or lubrication systems are either completely turned off or operating at a low capacity.

The central argument in this analysis is that the existence, frequency and consumption profile of those steady-state conditions are central to the assessment of energy efficiency on board. They are characterized by an operational environment where energy consumption is predictable, as it is dominated by steady-state consumption of various major consumers such as major pumps and big blowers. Especially when the vessel is idle, the crew on board has the opportunity to turn off - or at least operate at a low capacity- several systems such as the engine cooling and lubrication systems. The consumption of major consumers can be estimated and aggregated to estimate the expected fuel consumption. The expected fuel consumption allows comparison to the actual consumption of a particular vessel, and also across sister vessels that share the same design.

Other operations can be inherently difficult to assess and compare to each other, as they are often influenced by multiple factors that can be hard to predict. For example, when examining consumption patterns of the cargo systems during discharging, factors such as the cargo discharge rate and the backpressure from the storage tanks can vary a lot between terminals and result in large scatter in the data.

Lastly, there are certain patterns of operations that can indicate a lack of energy efficiency. One of those is the case where the vessels are standby at port. During standby the vessel is not turning off any equipment as it should be in a position to depart imminently. While it can be a request from terminals and port authorities to keep the vessel in standby as a safety precaution, when a vessel is systematically on standby it can indicate improper Finished-With-Engine (FWE) procedures either due to a lack of energy awareness from the crew or because the systems on-board cannot be operated efficiently.

The goal of the study is to assess the energy consumption patterns for a group of tanker vessels. The vessels are operating in the spot market (Stopford 2009), meaning that they trade without fixed schedule. The analysis focuses on two operational profiles: the first part focuses on cases when the vessels are at port and not sailing, and the second part when the vessel is under sea passage. Through discussions with onshore performance managers and marine engineers, six main vessel states were identified, as shown in Table 1.

Table 1: Operational profile breakdown and description of the identified operational conditions

Operational profile	Operating Condition	Description
Port stay analysis	Idle	The vessel is at port, and a series of systems can be safely turned off, or be operated at low capacity. No significant activities take place, meaning that the systems on board are operating in a steady state. Such operations can provide a basis for comparison using vessel-specific baselines.
	Static operations	The vessel is conducting operations while at port that require the use of various systems on board. Such operations may include for example tank cleanings, cargo heating and circulation, drifting, and cargo discharging. They can be seen as exceptional cases, where higher consumption is justified, and whose frequency and intensity varies depending on the trading profile of the vessel. Such operations can be difficult to compare even between vessels that share the same design.
	Improper Finished-With-Engine (FWE) procedures	Cases where excess equipment is run, that is not justified by the trading profile. Such equipment may include unnecessary parallel running of generators, sea water and fresh water pumps, fire pumps and hydraulic systems. Such operations can be justified in extraordinary circumstances, for example in the case of very short port stays, High Risk Areas (HRA) and drifting. Systematic presence may indicate a lack of shut-down procedures.

Sailing analysis	Steady sea passage	The vessel is sailing according instructions. No significant short-term operations take place, meaning that the systems on board are operating in a steady state and the vessel is sailing under stable speed. Such operations can provide a basis for comparison using vessel-specific baselines.
	Operations while sailing and slow steaming	The vessel is conducting operations while on sea passage, that require the use of various systems on board. Such operations may include for example tank cleanings, cargo heating and circulation, manoeuvrings in confined waters and deck operations. They can be seen as exceptional cases, where higher consumption is justified, and whose frequency and intensity varies depending on the trading profile of the vessel. Such operations can be difficult to compare even between vessels that share the same design.
	Excess equipment running	Cases where excess equipment is run, that is not justified by the trading profile. Such equipment may include unnecessary parallel running of generators, sea water and fresh water pumps, fire pumps, air compressors and hydraulic systems. Such operations can be justified in extraordinary circumstances, for example when transiting High Risk Areas (HRA). Systematic presence may indicate a lack of focus on energy efficiency on board, or system malfunctioning due to sub-par maintenance.

2.2 Data collection

To carry out this assessment, several data sources were combined. The primary tool for assessing energy consumption was noon reports. Noon reports refer to data collected every 24 hours at noon manually by the crew, and describe the operation of the ship over the last 24 hours. They are standard practice in many shipping companies and remain key tools for data collection (Poulsen & Johnson 2015). An alternative to noon reports are auto-logging systems. Auto logging systems rely on onboard sensors to collect data without manual intervention. And while such systems are increasingly seen as sources of value (Morlet et al. 2016), they suffer from bandwidth limitations as data needs to be transmitted via satellite. Therefore noon reports, despite their inherent limitations, are likely to continue as a prime data collection tool in the foreseeable future due to practical limitations with current auto logging systems.

Based on the noon reports, energy consumption data were divided per consumer and covered the auxiliary engines used for production of electricity, boilers, main engine, Inert Gas Generator and other minor consumers such as Framo pumps and incinerators. Operational data from the noon reports were used to assess the operational condition of the vessel. Such data included generator and oil fired boiler running hours, vessel's speed over ground and speed through water, weather conditions and sea water temperatures. Afterwards the data sources were consolidated into a single data set. The theoretical baselines for electricity consumption on board were determined based on information extracted from the vessel's equipment list, as stated in the newbuilding specifications. The consumption estimates for the equipment were validated using actual measurements on board.

Lastly, empirical data were included as well, and played a major role in the analysis. Positive and negative patterns of energy efficient operations were identified for a series of vessels and verified through discussions with senior officers and technical superintendents. During those groups selected performance patterns were analysed, evaluating them against the commercial schedule of the vessel. This validated data set provided the training set for supervised learning, and is discussed in Section 2.3.3.

2.3 Data analysis

2.3.1 Data preprocessing

With regards to data cleaning, missing values were dealt with by means of listwise deletion. So in cases were data was missing, the whole tuple was ignored(Han et al. 2012, p. 82). This was done in order to ensure maximum confidence in the data. It should be noted that listwise deletion did not result in massive losses of data, as missing values were present in less than 0.5% of the reports. Lastly, since noon reports are manually input in the system, boundaries on the minimum and maximum values were set to filter for clearly erroneous values.

2.3.2 Feature selection using Penalized Linear Discriminant Analysis

Feature selection is an important part of model building, and a necessity in many machine learning applications(Saeys et al. 2007). Especially in the presence of high dimensional data the inclusion of additional features leads to worse rather than better performance (Duda et al. 2001). Use and application of feature selection algorithms has multiple benefits, including reduced overfitting, faster and more cost-effective models and a deeper understanding into the underlying processes that generated the data. However, feature selection algorithms in classification problems add an extra layer of complexity, and their efficacy is often influenced by intrinsic properties of the data such as multimodality and the degree of overlap between classes (Saeys et al. 2007; Duda et al. 2001)

In this study, Penalized Linear Discriminant Analysis (PLDA) was employed as a screening tool to assess the discriminating abilities of each variable (Witten & Tibshirani 2011; Hastie et al. 1995). In our case, boundaries between the operating conditions are likely to be non-linear due to a mix of behavioural and technical constraints (see for example Myśków & Borkowski (2015) for the non-linear effect of slow steaming on oil fired boiler consumption). And as Linear Discriminant Analysis (LDA) can be too rigid in situations where class boundaries in predictor space are complex and non-linear (Hastie et al. 1995), we used a modified version of Penalized Linear Discriminant Analysis (PLDA) based on the work by Witten & Tibshirani (2011). The desired result is the value of the discriminant vector, which contains the values of the eigenvalues of the matrix product of the inverse of the within-group sums-of-squares and cross-product matrix and the between-groups sums-of-squares and cross-product matrix. The magnitudes of the eigenvalues are indicative of the features' discriminating abilities, and can be used to calculate the percentage of variance explained by that particular variable.

2.3.3 Classification using Multi Class Support Vector Machines

Support Vector Machines (SVMs) are multivariate artificial learning algorithms. SVMs rely on pre-processing the data and a non-linear mapping to separate data from two categories by a hyperplane (Duda et al. 2001), as shown in Figure 2. They can be used for supervised classification, as they can learn about group differences in a training set categorized by a priori knowledge and apply the model to assess new data points (Barber 2011). Support vector machines have been successfully used in a wide range of applications, including speech and image recognition (Burges 1998), fault detection in HVAC (Yan et al. 2014) , remaining useful life prediction (Sikorska et al. 2011), building energy consumption (Dong et al. 2005) and mental disease diagnosis (Koutsouleris et al. 2009).

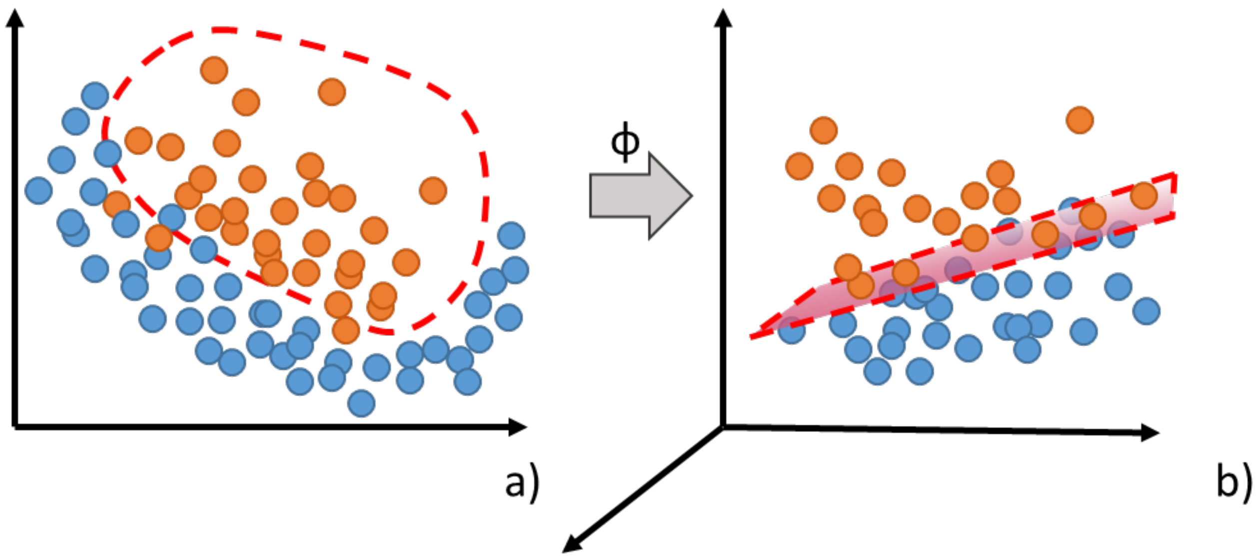


Figure 2: Schematic representation of support vector machine classification. a) A linear classifier cannot separate the two classes (illustrated as blue and orange) as the boundary (red dashed line) is non-linear b) A non-linear mapping (ϕ) maps the observations into a higher dimensional space

In this study we employed multi-class support vector machines used a radial basis functions kernel, which it facilitates the adaptive modelling of the interface between the classes and thus significantly improves classification performance. Implementation was based on the package “Kernlab“ in R (Karatzoglou et al. 2016).

To estimate the generalizability of the classification 5-fold cross-validation was performed (Zhang 1993). In 5-fold cross-validation the original sample is partitioned into five subsamples of equal size, and one subsample is used as a validation set for testing the model, while the other four are used for training the model. The process is repeated four times, so that all observations are used for both training and validation. The parameters C and gamma were determined through exhaustive grid search by minimizing the average validation error for those four runs. A flowchart of the proposed algorithm is shown in Figure 3.

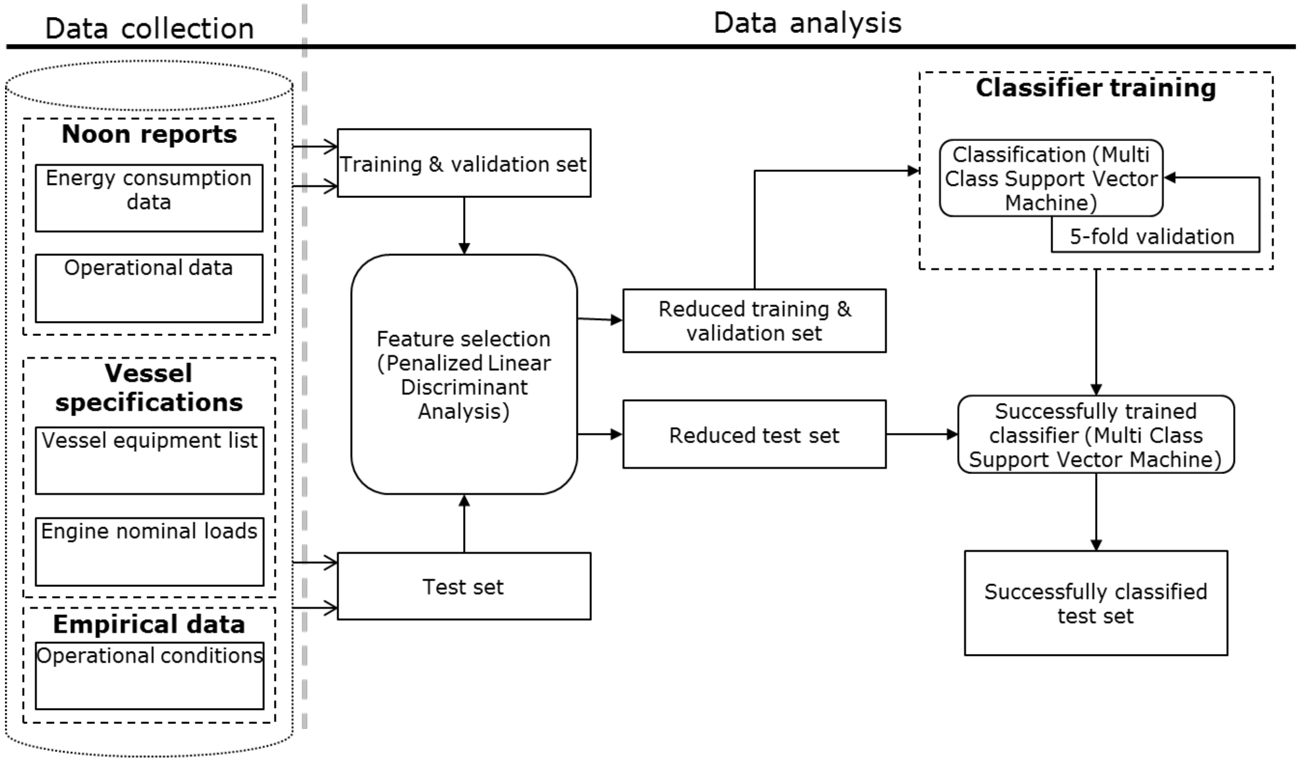


Figure 3: Flowchart of proposed algorithm

3 Results

3.1 Classifier performance

According to the Penalized Discriminant Analysis the five most important features, as ranked by the percentage of variance that each can explain are shown in Table 2. In total, features were selected so that at least 95% of the variance is retained in the reduced dataset.

Based on the results, one can make two interesting observations. First, the fact that generator running hours seems to be the most informative feature in both analyses. Furthermore, creating of additional features by combining existing features is likely to improve the performance of the algorithm, even in cases where features are highly correlated.

Table 2: Five most important features according to Penalized Discriminant Analysis, and their discriminative ability in regards to the percentage of variance explained

Operational profile	Feature name	Calculation process [Measuring units]	Percentage of variance explained by the feature
Port stay analysis	Normalized generator running hours	Total generator running hours / Noon report duration [%]	39%
	Normalized generator consumption against expected consumption	Fuel consumption per 24 hours/ Expected fuel consumption per 24 hours [%]	25%
	Inert Gas Generator & Framo Consumption	Fuel consumption per 24 hours in tons [tons]	13%
	Normalized Main engine consumption	Fuel consumption per 24 hours / Fuel consumption at Maximum Continuous Rating per 24 hours [%]	12%
	Oil Fired Boiler consumption	Oil Fired Boiler consumption per 24 hours [tons]	5%
Sailing analysis	Normalized generator	Total generator	41%

running hours	running hours / Noon report duration [%]	
Normalized generator consumption against expected consumption	Fuel consumption per 24 hours/ Expected fuel consumption per 24 hours [%]	16%
Logged Speed	Logged distance / Noon report duration [knots]	12%
Normalized generator consumption against maximum consumption	Fuel consumption per 24 hours / Fuel consumption at Maximum Continuous Rating per 24 hours[%]	10%
Inert Gas Generator & Framo Consumption	Fuel consumption per 24 hours in tons [tons]	10%

Table 3 shows the classification performance for the multi-class classifier. Precision, recall and the F-score was calculated for all operational conditions. The classifier shows high accuracy, supported by high F-score values. However, it proves somewhat less effective in identifying improper operating conditions for both operational profiles.

Table 3: Classification performance for the two operational profiles

	Port analysis			Sailing analysis		
	Idle	Static operations	Improper FWE procedures	Steady sea passage	Operations while sailing and slow steaming	Excess equipment running
Precision (%)	98%	99%	100%	98%	81%	100%
Recall (%)	99%	97%	80%	93%	94%	56%
F-score	0.99	0.98	0.89	0.96	0.87	0.72
Average accuracy (%)	98%			94%		

3.2 Classification performance for a group of tanker vessels

The external validity of the developed algorithm was examined by classifying the operational patterns of five test vessels. The vessels were evaluated for the same two-month period, and the results were manually checked and discussed with relevant stakeholders. Table 4 shows the classification results as a percentage of the time that vessels spend in each operational condition.

Table 4: Classification results

Vessel Name	Port analysis: Number of days at port	Idle	Static operations	Improper FWE procedures	Sailing analysis: Number of days sailing	Steady sea passage	Sailing operations and slow steaming	Excess equipment running
Vessel A	31	59%	41%	0 %	30	61%	39%	0%
Vessel B	31	44 %	50 %	6 %	30	100%	0%	0%
Vessel C	31	44 %	55 %	20 %	30	66%	33%	0%
Vessel D	32	18 %	61 %	21 %	29	71%	15%	15%
Vessel E	32	8 %	62 %	30 %	29	35%	4%	61%

Figure 4 shows the assessment results, in regards to the generator consumption during sailing and at port. Reports are excluded in cases where operations are present (see Table 1 for the argumentation against including operations in the evaluation). The analysis highlights the fact that differences in generator consumption can be traced down to the way the systems are operated in practice.

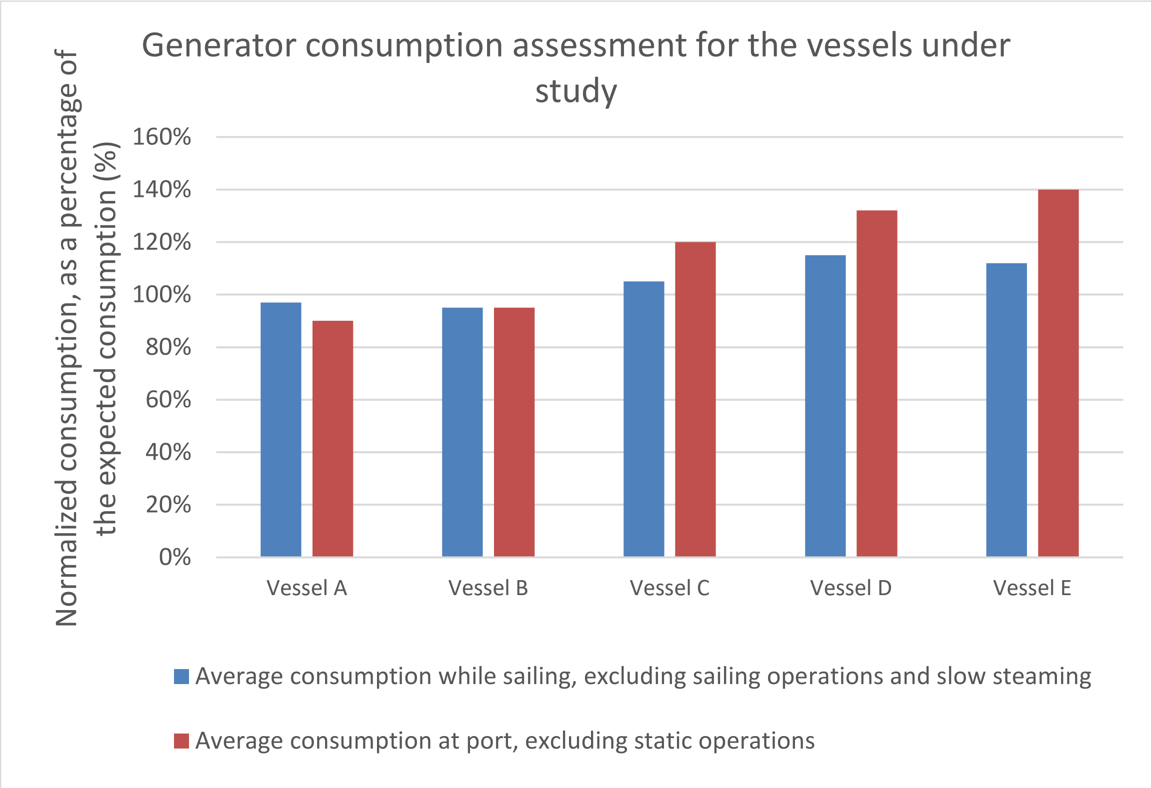


Figure 4: Assessment of average generator consumption for the five vessels under study

The results of the analysis highlighted the significant differences between the five vessels, and triggered a deeper investigation for Vessel E into the root causes behind the discrepancies. The investigation revealed that the efficiency gap was due to a mix of technical and behavioural causes. Addressing those issues resulted in yearly savings of approximately USD 50,000 for the average fuel prices in 2015.

4 Conclusions

This article describes a methodology for identifying operational patterns in regards to the power management on-board. To the best of our knowledge, this study is the first to evaluate the efficacy of machine learning algorithms within energy management in shipping. The proposed methodology is conceptually simple, and able to deal with multiple data sources. It employs established tools, and exhibits high prediction accuracy and low misclassification rates. At this point, it should be noted that similar results could be obtained using other machine learning algorithms such as neural networks or a more structured algorithm like the one described in (Trodden et al. 2015). Nevertheless, the non-linear character of the data together with the flexibility of Multi-Class Support Vector Machines supported their choice in the context of the study.

Regarding the managerial implications of the study, the results show that focus on power management on board can vary widely among vessels. Most importantly, identifying these differences and alleviating their root causes can lead to a sustained reduction in life cycle costs. Future work could focus on applying the same methodology on other areas within performance management, including hull and propeller performance and steam production on board. Moreover, future work could integrate more measurements -including individual equipment running hours- and expand to evaluate data streams from auto logging systems

Acknowledgement

The authors would like to thank the TORM foundation for its financial support

References

Barber, D., 2011. *Bayesian Reasoning and Machine Learning*,

Burges, C.J.C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2), pp.121–167.

Chalal, M.L. et al., 2016. Energy planning and forecasting approaches for supporting physical improvement strategies in the building sector: A review. *Renewable and Sustainable Energy Reviews*, 64, pp.761–776.

Cichowicz, J., Theotokatos, G. & Vassalos, D., 2015. Dynamic energy modelling for ship life-cycle performance assessment. *Ocean Engineering*, 110, pp.49–61.

Court, D., 2015. Getting big impact from big data. *McKinsey Quarterly*, (1), pp.52–60. Available at:

http://www.mckinsey.com/business-functions/business-technology/our-insights/getting-big-impact-from-big-data [Accessed April 23, 2016].

DNV - GL, 2016. *A New Reality: the outlook for the oil and gas industry 2016*,

Dong, B., Cao, C. & Lee, S.E., 2005. Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, 37(5), pp.545–553.

Duda, R.O., Hart, P.E. & Stork, D.G., 2001. *Pattern classification*,

Han, J., 2012. *Data mining : concepts and techniques*, Elsevier/Morgan Kaufmann.

Hastie, T., Buja, A. & Tibshirani, R., 1995. Penalized Discriminant Analysis. *The Annals of Statistics*, 23(1), pp.73–102.

He, H., Zhang, J. & Li, G., 2016. Model Predictive Control for Energy Management of a Plug-in Hybrid Electric Bus. *Energy Procedia*, 88, pp.901–907.

Karatzoglou, A., Smola, A. & Hornik, K., 2016. Kernel-Based Machine Learning Lab [R package kernlab version 0.9-24].

Koutsouleris, N. et al., 2009. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Archives of general psychiatry*, 66(7), pp.700–12.

Lee, D. & Cheng, C.-C., 2016. Energy savings by energy management systems: A review. *Renewable and Sustainable Energy Reviews*, 56, pp.760–777.

Morlet, A. et al., 2016. *Intelligent Assets: Unlocking the Circular Economy Potential*,

Murray, D.M. et al., 2016. Understanding usage patterns of electric kettle and energy saving potential. *Applied Energy*, 171, pp.231–242.

Myśków, J. & Borkowski, T., 2015. Emission Inventory of Exhaust Gases From Oil-Fired Ship Boilers. *Journal of KONES. Powertrain and Transport*, 19(4), pp.487–494.

NATO Shipping Centre, 2016. Skip navigation links Updated "Guidelines for Owners, Operators and Masters for protection against piracy in the Gulf of Guinea region."

Poulsen, R.T. & Johnson, H., 2015. The logic of business vs. the logic of energy management practice: Understanding the choices and effects of energy consumption monitoring systems in shipping companies. *Journal of Cleaner Production*.

Poulsen, R.T. & Sornn-Friese, H., 2015. Achieving energy efficient ship operations under third party management: How do ship management models influence energy efficiency? *Research in Transportation Business and Management*, 17, pp.41–52.

Saeys, Y., Inza, I. & Larrañaga, P., 2007. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19), pp.2507–2517.

Sapankevych, N. & Sankar, R., 2009. Time Series Prediction Using Support Vector Machines: A Survey. *IEEE Computational Intelligence Magazine*, 4(2), pp.24–38.

Sikorska, J.Z., Hodkiewicz, M. & Ma, L., 2011. Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*, pp.1803–1836.

Stopford, M., 2009. *Maritime Economics* 3rd ed., Routledge.

Thompson, D.N., 2012. *Oracles : how prediction markets turn employees into visionaries*, Harvard Business Review Press.

Trodden, D.G. et al., 2015. Fuel usage data analysis for efficient shipping operations. *Ocean Engineering*, 110, pp.75–84.

Westerman, G. et al., 2011. Digital Transformation: A Road-Map for Billion-Dollar Organizations. *MIT Center for Digital Business and Capgemini Consulting*, pp.1–68.

Witten, D.M. & Tibshirani, R., 2011. Penalized classification using Fisher's linear discriminant. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 73(5), pp.753–772.

Yan, K. et al., 2014. ARX model based fault detection and diagnosis for chillers using support vector machines. *Energy and Buildings*, 81, pp.287–295.

Zhang, P., 1993. Model Selection Via Multifold Cross Validation. *The Annals of Statistics*, 21(1), pp.299–313.